CS 6375
ASSIGNMENT 2

Names of students in your group:
Yagna Srinivasa Harsha Annadata (YXA210024)
Vishruth Reddy Chinthi Reddy (VXC220020)

Number of free late days used: 2

Note: You are allowed a **total** of 4 free late days for the **entire semester**. You can use at most 2 for each assignment. After that, there will be a penalty of 10% for each late day.

Please list clearly all the sources/references that you have used in this assignment.

1. Neural Networks - GeeksForGeeks
2. https://www.tensorflow.org/api_docs/python/tf/keras

Part2 – Programming Part

Google Colab link:
https://colab.research.google.com/drive/1CEuizREFsN2JXqj2BCvgGxUfjMMDguF2?usp=sharing

Data Link:
Dataset Link:
https://raw.githubusercontent.com/YagnaAnnadata/Projects/main/LinearRegression/Dataset/winequality/winequality-red.csv
Original Dataset Link: https://archive.ics.uci.edu/dataset/186/wine+quality

Data Set Used: Wine Quality

Quality of wine is the most important feature in wine making. The goal is to model wine quality based on physicochemical tests. The features which determine the quality of wine here are Fixed acidity, volatile acidity, citric acid, residual sugar, chlorides, free sulfur dioxide, total sulfur dioxide, density, pH, sulphates, alcohol.

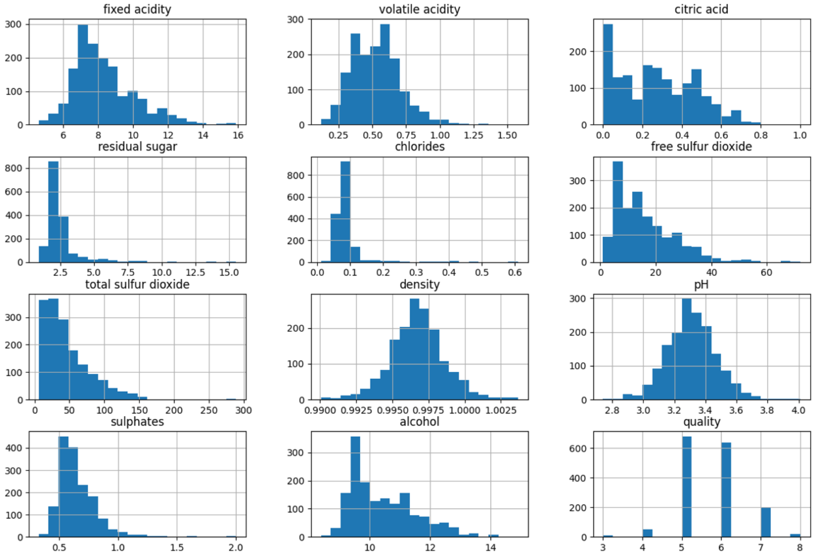
Variables List:

Independent Variables:
x1 = fixed acidity
x2 = volatile acidity
x3 = citric acid
x4 = residual sugar
x5 = chlorides
x6 = free sulfur dioxide
x7 = total sulfur dioxide
x8 = density
x9 = pH
x10 = sulphates
x11 = alcohol

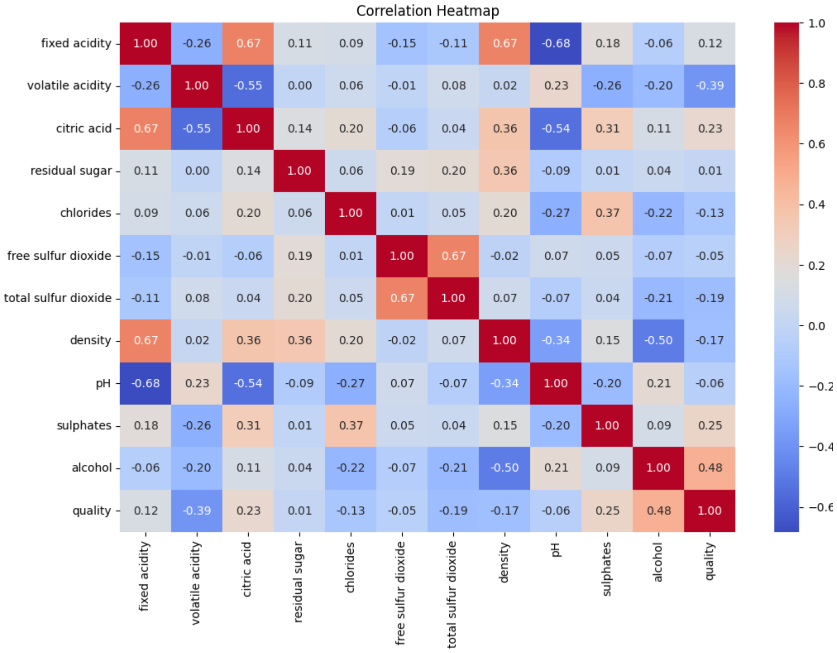
Dependent Variables:
Y = Quality

Distribution of variables:

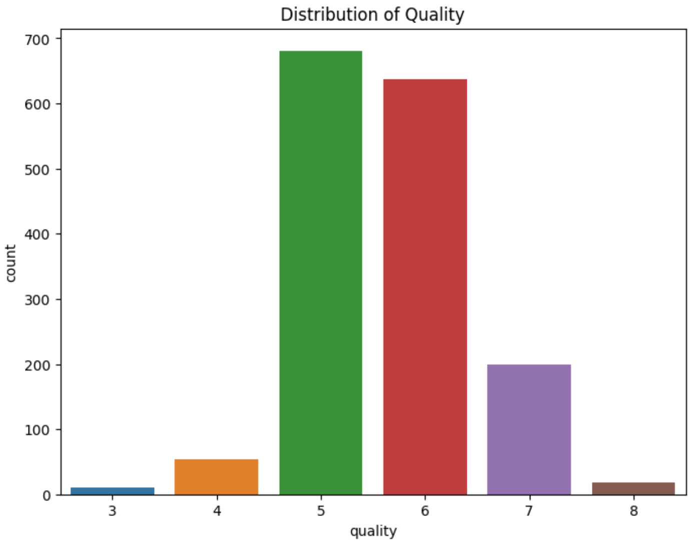
The below graphs show the histogram of the variables.



Correlation Matrix:



	fixed acidity	volatile acidity	citric acid	residual sugar	chlorides	free sulfur dioxide	total sulfur dioxide	density	pH	sulphates	alcohol	quality
fixed acidity	1.00	-0.26	0.67	0.11	0.09	-0.15	-0.11	0.67	-0.68	0.18	-0.06	0.12
volatile acidity	-0.26	1.00	-0.55	0.00	0.06	-0.01	0.08	0.02	0.23	-0.26	-0.20	-0.39
citric acid	0.67	-0.55	1.00	0.14	0.20	-0.06	0.04	0.36	-0.54	0.31	0.11	0.23
residual sugar	0.11	0.00	0.14	1.00	0.06	0.19	0.20	0.36	-0.09	0.01	0.04	0.01
chlorides	0.09	0.06	0.20	0.06	1.00	0.01	0.05	0.20	-0.27	0.37	-0.22	-0.13
free sulfur dioxide	-0.15	-0.01	-0.06	0.19	0.01	1.00	0.67	-0.02	0.07	0.05	-0.07	-0.05
total sulfur dioxide	-0.11	0.08	0.04	0.20	0.05	0.67	1.00	0.07	-0.07	0.04	-0.21	-0.19
density	0.67	0.02	0.36	0.36	0.20	-0.02	0.07	1.00	-0.34	0.15	-0.50	-0.17
pH	-0.68	0.23	-0.54	-0.09	-0.27	0.07	-0.07	-0.34	1.00	-0.20	0.21	-0.06
sulphates	0.18	-0.26	0.31	0.01	0.37	0.05	0.04	0.15	-0.20	1.00	0.09	0.25
alcohol	-0.06	-0.20	0.11	0.04	-0.22	-0.07	-0.21	-0.50	0.21	0.09	1.00	0.48
quality	0.12	-0.39	0.23	0.01	-0.13	-0.05	-0.19	-0.17	-0.06	0.25	0.48	1.00

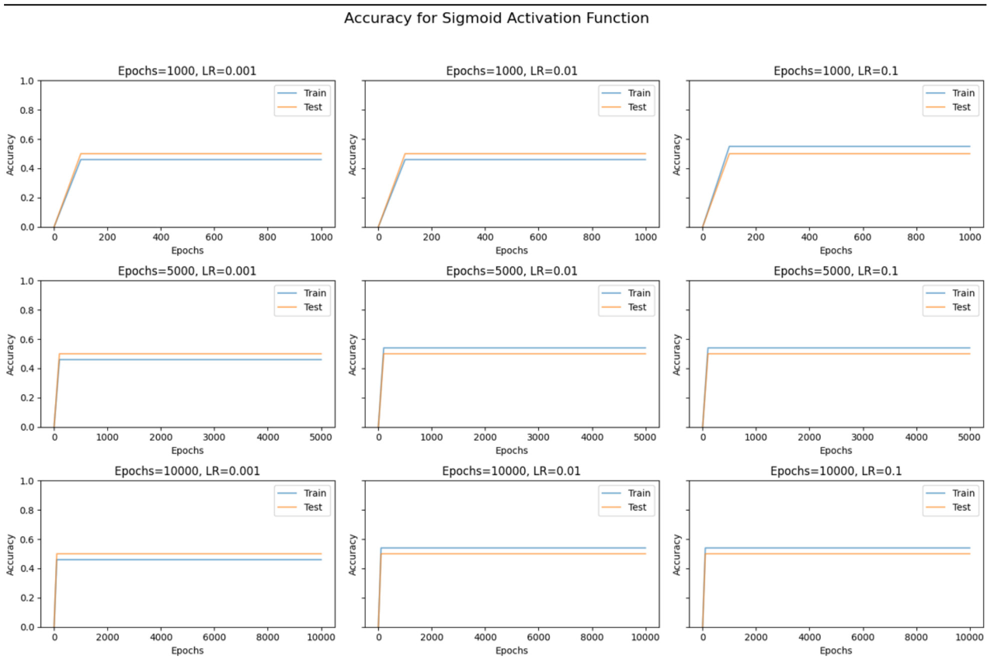


Model Summary:

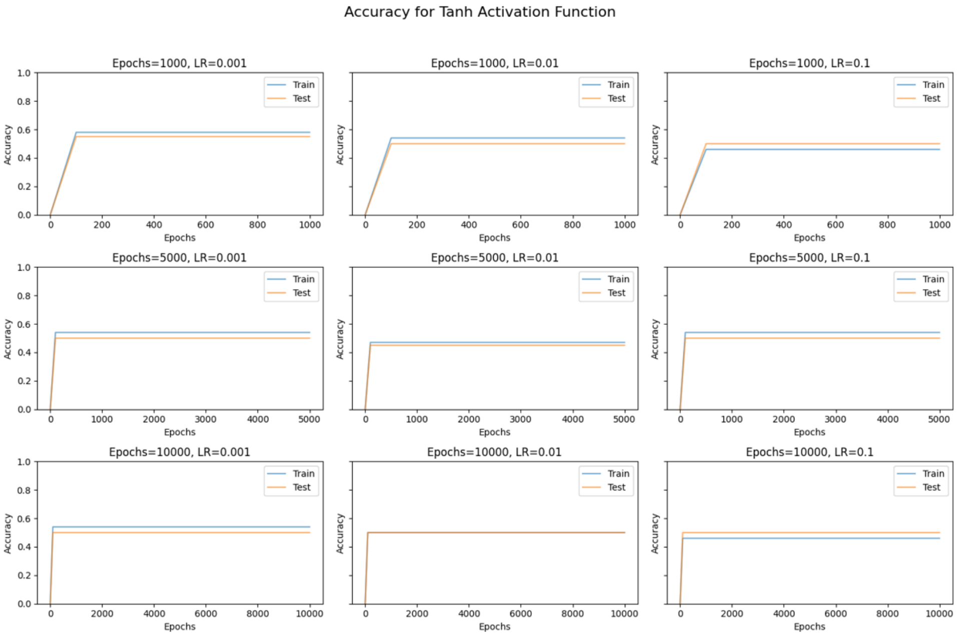
Activation Function	Learning rates	0.001	0.01	0.1	0.001	0.01	0.1	0.001	0.01	0.1
	Epoch	1000			5000			10000		
Sigmoid	Train Accuracy	0.46	0.46	0.55	0.46	0.54	0.54	0.46	0.54	0.54
	Test Accuracy	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50
Tanh	Train Accuracy	0.58	0.54	0.46	0.54	0.47	0.54	0.54	0.50	0.46
	Test Accuracy	0.55	0.50	0.50	0.50	0.45	0.50	0.50	0.50	0.50
Relu	Train Accuracy	0.56	0.54	0.54	0.49	0.54	0.54	0.45	0.54	0.54
	Test Accuracy	0.55	0.50	0.50	0.70	0.50	0.50	0.80	0.50	0.50

Plots:

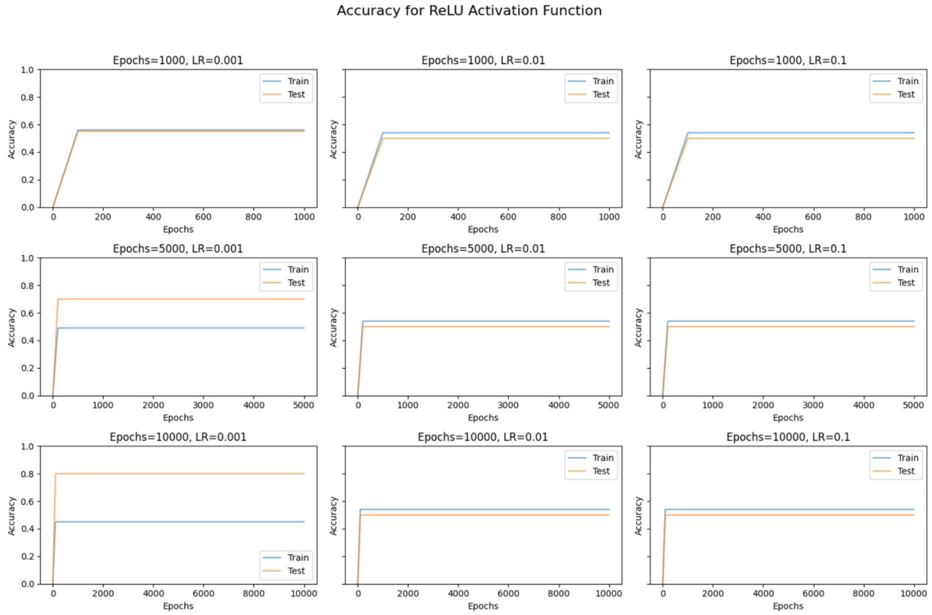
Training and Test accuracy with different Learning Rates and Epochs for Sigmoid Function:



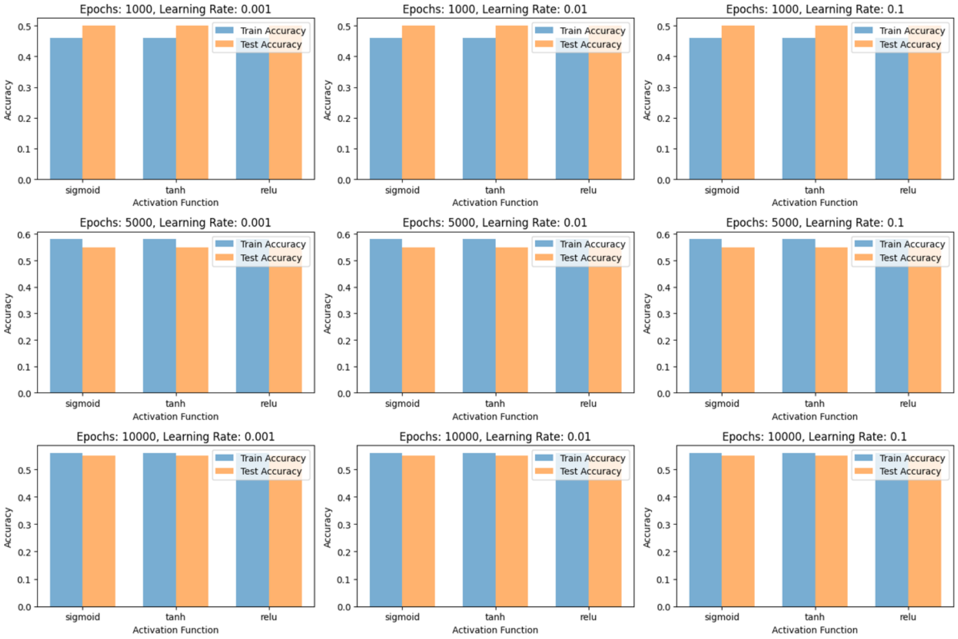
Training and Test accuracy with different Learning Rates and Epochs for Tanh Function:



Training and Test accuracy with different Learning Rates and Epochs for Relu Function:



Comparisons of Training and Test accuracy with different Learning Rates and Epochs for Activation Functions:



Libraries Used:

- Numpy
- Pandas
- Scikit-learn
- Matplotlib
- Seaborn
- Keras

Conclusion:

As observed from the above table and plots, we find that Relu Activation function performs the best. Here the Relu activation function, the graph converges comparatively faster because of its functionality of choosing either zero and x.

The following is observed from the summary table:
This case displays higher Accuracy compared to all
other results.
Activation Function: Relu, Learning Rate: 0.001, Epochs: 10000
Train Accuracy: **0.4500**, Test Accuracy: **0.8000**